刘仕国 . 新质生产力对社会科学数据应用方法的影响［J］. 文献与数据学报，2025，7（1）：011-013.

新质生产力对社会科学数据应用方法的影响

刘仕国

（中国社会科学院世界经济与政治研究所，北京　100732）

DOI：10.31193/SSAP.J.ISSN.2096-6695.2025.01.04

在实证主义研究范式中，数据是验证理论和发现规律的关键。社会科学研究对数据的重视程度逐渐提高，数据应用方法逐渐得到发展。在此过程中，生产力起到了至关重要的作用。19 世纪中叶以前，社会科学研究主要依赖传统的思辨和演绎方法，数据应用有限。20 世纪 50 年代左右，行为主义革命提高了社会科学对数据的重视，政治学、社会学和教育学等学科开始强调数据的收集与整理，使用统计学、形式模型和计算机模拟等开展研究，量化分析开始兴起。20 世纪后半叶，计算机和信息技术的发展，推动研究者开始使用海量数据开展应用研究，促进了定量社会科学方法的广泛应用。

当前，新质生产力体现了当今世界新一轮科技革命和产业变革的大趋势、大潮流，正在并将继续深刻影响社会科学数据应用方法的发展。

其一，极大拓展了社会科学数据的来源、类型和形式，推动社会科学研究从传统的小样本、静态、结构化数据向多源、动态、多模态数据转变。

数据来源从“有限渠道”拓展到“全域覆盖”。传统数据主要来自问卷调查、官方统计、文献档案等。新质生产力通过技术手段使数据来源更加多元化和实时化，带来社交媒体与网络行为数据、物联网与传感器数据、在线平台与数字足迹、区块链与分布式账本等。

推动社会科学数据类型从单一文本、数值型数据扩展到多模态、非结构化数据，更贴近真实社会场景。传统数据以结构化数据为主，如统计表格、问卷量表、档案记录。新技术催生出非结构化文本数据、图像与视频数据、时空轨迹数据、生物与生理数据以及虚拟空间数据等。

推动数据形式从静态的“快照”升级为动态、实时、可交互的“流数据”。通过 5G 和边缘计算，实时获取社交媒体舆情波动、金融市场情绪变化。例如，用户与人工智能聊天机器人的对话成为人机协作生成型数据，生成式人工智能模拟社会实验生成合成数据与仿真数据。

其二，提高数据全生命周期管理能力，为社会科学数据应用方法奠定坚实数据基础。

感知与连接技术提高了数据获取能力。传感器、智能终端等物联网硬件设备实时采集物理世界数据。5G/6G 通信技术支持海量设备互联，实现高密度数据实时传输。卫星遥感与无人机提供了覆盖全球的地理空间数据获取能力。边缘计算技术在数据源头进行初步处理，减少数据传输压力。通过计算机视觉、语音识别等人工智能驱动的智能感知技术自动提取结构化数据。

分布式与自动化技术提供了强大的数据收集能力。云计算平台等弹性资源池支持大规模数据

［作者简介］刘仕国，男，研究员，研究方向为世界经济统计学和开放宏观计量经济学，Email：liusg@cass.org.cn。

汇聚。分布式架构等边缘节点实现本地化数据缓存与预处理。流数据处理技术实时收集并处理动态数据流。自动化爬虫与应用程序接口自动抓取网络公开数据。

智能处理与标准化提升了数据整理能力。高性能计算集群快速清洗、标注和转换异构数据。人工智能芯片加速数据预处理，如图像去噪、文本分词等。自然语言处理技术自动分类、去重、实体识别。数据湖与数据中台统一存储格式，实现跨源数据融合。

数据存储技术确保数据高容量与高可用性。分布式存储系统支持 PB 级数据低成本存储。量子存储等新型存储介质突破传统容量限制。区块链技术确保数据不可篡改。通过人工智能算法，实行冷热数据分层管理，从而优化存储资源分配。

数据传播实现数据的安全共享与高效分发。内容分发网络布局全球节点，提高数据访问速度。量子通信网络保障数据传输的绝对安全。联邦学习与隐私计算技术实现数据“可用不可见”，如医疗数据跨机构共享。智能合约与数据市场技术使数据交易通过区块链平台实现交易自动化。

其三，深度重塑了社会科学数据应用的方法论，推动社会科学研究从“小样本归纳”向“全样本计算”、从“静态分析”向“动态推演”、从“人工解释”向“智能决策”转变，甚至引发从“理论驱动”到“数据驱动”的范式变革。

推动研究流程从线性到闭环迭代。传统流程一般为“假设—数据收集—分析—验证”。新质生产力驱动研究流程创新。“实时数据反馈闭环”：物联网传感器与边缘计算实现“数据采集—分析—策略调整”实时联动。人工智能辅助假设生成：利用自然语言处理挖掘文献库生成研究假设。仿真验证前置：通过数字孪生技术构建虚拟社会系统，预演政策效果。例如，联合国开发计划署使用“卫星影像 + 人工智能算法”实时监测贫困区域设施变化，动态调整援助策略。

推动分析工具从统计软件到智能引擎的升级。传统分析工具依赖 SPSS、Stata 等统计软件处理结构化数据，难以应对非结构化、高维数据，而新技术则实现了如下赋能。多模态数据融合分析：知识图谱整合人物关系、事件时空轨迹与文本语义，如反腐研究中利益网络可视化。复杂性科学工具：应用基于机构的建模模拟群体行为涌现，如疫情中谣言传播的多智能体仿真。因果推断革命：利用强化学习构建反事实推理模型，如评估教育扶贫政策的长期因果效应。

实现从实验室到“社会实验室”的革新。传统试验方法的瓶颈在于田野实验成本高、控制变量难，实验室环境脱离真实社会场景，而新质生产力则带来了全域社会实验场、混合现实实验、合成数据实验等。例如，斯坦福大学（Stanford University）在《第二人生》（Second Life）虚拟世界中复刻现实城市，研究灾害应急响应中的集体行动逻辑。

推动协作模式从封闭向开放转型。传统科学研究协作存在数据孤岛、方法论不透明等壁垒，新质生产力则催生了如下新的协作生态。区块链存证与溯源：确保研究过程不可篡改开源社区共建，如构建社会科学模型库共享预训练模型。众包数据标注：如亚马逊机械杜克（Arnazon Mechanical Turk）结合人工智能自动质检，低成本完成多语言社会调查标注。成功案例之一就是全球气候行动追踪项目整合卫星数据、企业碳排放报告与非政府组织调查，通过区块链确权实现跨国联合分析。

决策支持从经验判断到数据智能跃迁。传统决策依赖专家经验与滞后统计指标，而智能决策系统基于广域信息系统和人工智能模型。例如，社会风险预警系统应用自然语言处理和时间序列

分析预测社会风险，政策模拟推演引擎基于多智能体建模评估最低工资调整对产业链影响。

应该看到，新质生产力对社会科学数据应用方法的影响目前处于早期阶段，更深刻、更广泛的变革正在到来。

立足实践探索，完善数据资产价格形成机制

王建冬[1]　孙　静[2]

（1. 国家发展改革委价格监测中心，北京　100837；2. 北京大学工学院，北京　100871）

DOI：10.31193/SSAP.J.ISSN.2096-6695.2025.01.05

数字经济时代，数据成为关键生产要素，建立健全数据资产定价机制可为促进数据要素流通、盘活数据资产、释放数据价值提供有力支撑。当前，数据资产“千人千价”的现象较为普遍，还需支持市场实践，强化政策引导，推动数据要素市场体系建设，健全数据资产市场价格形成机制。

1　数据资产价格形成机制尚处于探索期

数据要素是我国首先提出的重大理论创新，数据资产定价是全球性难题之一，在世界范围内没有成熟经验，需持续探索。

1.1　数据资产定价为激活数据要素潜能提供支撑

目前，社会各行各业都积累了大量“沉睡”的数据资产。随着数据资源入表等实践的推进，对企业拥有的数据资源进行评估定价并确认为资产的工作提上日程。新质生产力以数据要素的深度应用为核心，依托数字化技术优化生产流程、创新商业模式，而数据资产的市场化定价正是释放这一潜能的关键前提。探索数据资产市场定价，一方面能够促使数据供给主体对数据资产价值形成更为深刻的认识，充分调动其积极性主动性，强化数据资产保护。如此一来，处于“沉睡”状态的数据资产将被盘活，这有助于达成数据资产保值增值的目标，提升数据供给的质量。另一方面，科学合理的数据资产市场定价也将促进数据要素市场合规流通，提高数据商和第三方专业服务机构专业化运营能力，有助于构建规范高效的数据交易场所，优化数据资源配置。

1.2　数据资产市场定价的基本原则逐步确立

党的十九届四中全会正式将数据列为生产要素以来，随后《关于构建更加完善的要素市场化

［作者简介］王建冬，男，研究员，研究方向为数据定价、大数据分析、数据要素治理，Email: wangjd_wyzx@126.com；孙静，女，助理馆员，研究方向为数据分析、科学评价等，Email：sunjingcoe@pku.edu.cn。